



Krylov integrators for Hamiltonian systems

Timo Eirola¹ · Antti Koskela²

Received: 14 December 2017 / Accepted: 1 October 2018 / Published online: 10 October 2018
© The Author(s) 2018

Abstract

We consider Arnoldi-like processes to obtain symplectic subspaces for Hamiltonian systems. Large dimensional systems are locally approximated by ones living in low dimensional subspaces, and we especially consider Krylov subspaces and some of their extensions. These subspaces can be utilized in two ways: by solving numerically local small dimensional systems and then mapping back to the large dimension, or by using them for the approximation of necessary functions in exponential integrators applied to large dimensional systems. In the former case one can expect an excellent energy preservation and in the latter this is so for linear systems. We consider second order exponential integrators which solve linear systems exactly and for which these two approaches are in a certain sense equivalent. We also consider the time symmetry preservation properties of the integrators. In numerical experiments these methods combined with symplectic subspaces show promising behavior also when applied to nonlinear Hamiltonian problems.

Keywords Hamiltonian systems · Exponential integrators · Krylov subspace methods · Symplectic integrators · Symmetric integrators · Hamiltonian Lanczos algorithm

Mathematics Subject Classification 65P10 · 37M15 · 65F60 · 65F10

1 Introduction

Symplectic methods have shown to be very effective in long time integration of Hamiltonian systems (see [11]). Many of them are implicit and necessitate the solution of

In memory of Timo Eirola (1951–2016).

Communicated by Christian Lubich.

✉ Antti Koskela
antti.h.koskela@helsinki.fi

¹ Department of Mathematics and Systems Analysis, Aalto University, Aalto, Finland

² Department of Mathematics and Statistics, University of Helsinki, Helsinki, Finland

systems of equations. If the differential equation system is large and sparse, a natural approach is to use Krylov subspace techniques to approximate the solution of the algebraic equations. A related approach is to use Krylov approximations of the matrix exponential in exponential integrators (see [13]). This has turned out to be a superior technique for several classes of problems.

Krylov subspace techniques can be viewed as local low dimensional approximations of the large system. For Hamiltonian systems standard Arnoldi-type iterations produce low dimensional systems that are no longer Hamiltonian. In this paper special attention is paid to produce symplectic bases for Krylov subspaces.

This is an extended version of the slides by Eirola, presented at the Workshop on Exponential Integrators in Innsbruck in 2004 (see [7]). Some of the results were introduced in the master's thesis of the second author [15] which was supervised by Eirola. The original ideas of Eirola came from considering linear Hamiltonian systems in \mathbb{R}^{2n} as \mathbb{R} -linear systems in \mathbb{C}^n (see [8]). The slides [7] were written using that language, but the present version is written in a more standard form.

2 Hamiltonian systems

Given a smooth function $H : \mathbb{R}^{2n} \rightarrow \mathbb{R}$, consider the Hamiltonian system

$$\mathbf{x}'(t) = \mathbf{J}_n^{-1} \nabla H(\mathbf{x}(t)), \quad \mathbf{x}(0) = \mathbf{x}_0, \quad (2.1)$$

where $\mathbf{J}_n = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \in \mathbb{R}^{2n \times 2n}$. A matrix $\mathbf{A} \in \mathbb{R}^{2n \times 2n}$ is called Hamiltonian, if $\mathbf{A}^T = \mathbf{J}_n \mathbf{A}$, and symplectic, if $\mathbf{A}^T \mathbf{J}_n \mathbf{A} = \mathbf{J}_n$. The Jacobian of $\nabla H(\mathbf{x})$ is a symmetric matrix at every point. Thus $D\mathbf{J}_n^{-1} \nabla H(\mathbf{x})$ is a Hamiltonian matrix.

The system (2.1) has a unique solution and we write it $\mathbf{x}(t) = \boldsymbol{\varphi}^t(\mathbf{x}_0)$. It holds:

- Energy is preserved: $H(\mathbf{x}(t))$ is constant in t .
- For every t the mapping $\boldsymbol{\varphi}^t : \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$ is symplectic (or canonical), that is, its derivative is a symplectic matrix at every point.
- The mapping $\boldsymbol{\varphi}^t$ is time symmetric, i.e. $\boldsymbol{\varphi}^{-t}(\mathbf{x}(t)) = \mathbf{x}_0$ for every t .

Symplectic integrators produce symplectic one step maps for Hamiltonian systems (see [11]). For example, the implicit midpoint rule

$$\mathbf{x}_{j+1} = \mathbf{x}_j + h \mathbf{J}_n^{-1} \nabla H((\mathbf{x}_j + \mathbf{x}_{j+1})/2)$$

is such. For linear systems, i.e., when H is of the form $H(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{S} \mathbf{x} + \mathbf{c}^T \mathbf{x}$, \mathbf{S} symmetric, the energy is also preserved in the numerical solution with this and many other symplectic methods. One step methods are called symmetric the map given by the integrator is time symmetric, i.e. changing h to $-h$ is equivalent to switching \mathbf{x}_j and \mathbf{x}_{j+1} . The implicit midpoint rule, for example, is symmetric.

For large systems implicit methods may become expensive. In this paper we consider several low dimensional Hamiltonian approximations and the use of exponential integrators for these.

3 Symplectic subspaces and low dimensional approximations

Recall some basic definitions and properties in \mathbb{R}^{2n} (see e.g. [1]). Denote the non-degenerate skew-symmetric bilinear form $\omega(x, y) := x^T J_n y$. A subspace V is *isotropic*, if $\omega(x, y) = 0$ for all $x, y \in V$, and a subspace W is *symplectic*, if for every nonzero $x \in W$ there exists $y \in W$ such that $\omega(x, y) \neq 0$. Then the dimension of W is even. A basis $e_1, \dots, e_k, f_1, \dots, f_k \in W$ is called symplectic, or a *Darboux basis*, if for all $i, j = 1, \dots, k$ holds $\omega(e_i, e_j) = \omega(f_i, f_j) = 0$ and $\omega(e_i, f_j) = \delta_{i,j}$.

If V is an isotropic subspace with an orthonormal basis e_1, \dots, e_k , then e_1, \dots, e_k are also ω -orthogonal and $W = V \oplus J_n V$ is a symplectic subspace and $e_1, \dots, e_k, J_n^{-1}e_1, \dots, J_n^{-1}e_k$ is a symplectic basis of W .

We call also a matrix $U \in \mathbb{R}^{2n \times 2k}$ symplectic, if $U^T J_n U = J_k$, where $J_k = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \in \mathbb{R}^{2k \times 2k}$. Then $U^\dagger = J_k^{-1} U^T J_n$ is a left inverse of U if and only if U is symplectic.

We will consider local approximations of the Hamiltonian system

$$x'(t) = f(x(t)) = J_n^{-1} \nabla H(x(t)). \tag{3.1}$$

Assume that at a point $x_0 \in \mathbb{R}^{2n}$ we are given a symplectic matrix $U \in \mathbb{R}^{2n \times 2k}$. Consider the Hamiltonian system in \mathbb{R}^{2k} corresponding to the function $\eta(\xi) = H(x_0 + U\xi)$. Then we get

$$\xi' = J_k^{-1} \nabla \eta(\xi) = J_k^{-1} U^T \nabla H(x_0 + U\xi),$$

which is Hamiltonian in \mathbb{R}^{2k} . Denote $U^\dagger = J_k^{-1} U^T J_n$. Then

$$\xi'(t) = U^\dagger f(x_0 + U\xi(t)). \tag{3.2}$$

One strategy is to solve (3.2) numerically from $\xi_0 = 0$ up to $\xi_1 \approx \xi(t_1)$ and set $x_1 = x_0 + U\xi_1$. Clearly, if we use an energy preserving scheme for the system (3.2), we will conserve the energy of the large system too, i.e. $H(x_1) = H(x_0)$.

Note that if the sets of constant energy of the original system are bounded, then they are such for the small dimensional approximations too. This implies that the approximations inherit stability of equilibria in a natural way.

We will consider also another strategy: instead of solving low-dimensional systems, we approximate the matrix functions appearing in exponential integrators in the low dimensional space \mathbb{R}^{2k} .

The idea of approximating a Hamiltonian system by another of smaller dimension is not new. See, for example the discussion in [16]. A novelty here is to use local (later Krylov subspace) approximations.

If U is symplectic and does not depend on x_0 , then using a symplectic method for (3.2) induces a map $\psi : x_0 \rightarrow x_1$ that is symplectic in $R(U)$, that is

$$\omega(D\psi(x_0)d, D\psi(x_0)\tilde{d}) = \omega(d, \tilde{d}) \quad \text{for all } d, \tilde{d} \in R(U).$$

But in order to get efficient algorithms we let U to depend on \mathbf{x}_0 and then this approach generally does not produce a symplectic map.

4 Exponential integrators

The use of approximations of the matrix exponential as a part of time propagation methods for differential equations has turned out to be very effective (see e.g. [13]). We consider application of three second order exponential integrators to the Hamiltonian system (2.1). In what follows, the matrix \mathbf{H} denotes the Jacobian of the right hand side of (3.1) at \mathbf{x}_0 , i.e., $\mathbf{H} = D\mathbf{f}(\mathbf{x}_0)$. The methods can be seen as exponential integrators applied to semilinear equations which are local linearizations of (2.1). In the literature methods of this type are also called *exponential Rosenbrock methods* [14].

4.1 Exponential Euler method

As a first method we consider the exponential Euler method (EE)¹

$$\mathbf{x}_+ = \mathbf{x} + h \phi(h\mathbf{H})\mathbf{f}(\mathbf{x}), \quad (4.1)$$

where $\phi(z) = \int_0^1 e^{tz} dt = z^{-1}(e^z - 1)$ and $\mathbf{H} = D\mathbf{f}(\mathbf{x})$.

Note that if \mathbf{f} is linear, then $\mathbf{x}_+ = \varphi^h(\mathbf{x})$, i.e., the method gives exact values of the solution.

Assume now that the system $\mathbf{x}' = \mathbf{f}(\mathbf{x})$ is Hamiltonian in \mathbb{R}^{2n} and $\mathbf{U} \in \mathbb{R}^{2n \times 2k}$ is symplectic. Then $\mathbf{H} = D\mathbf{f}(\mathbf{x})$ is a Hamiltonian matrix as well as $\mathbf{F} = \mathbf{U}^\dagger \mathbf{H} \mathbf{U}$.

If we use the exponential Euler method (4.1) for the low dimensional system (3.2) we produce

$$\mathbf{x}_+ = \mathbf{x} + h \mathbf{U} \phi(h\mathbf{F}) \mathbf{U}^\dagger \mathbf{f}(\mathbf{x}). \quad (4.2)$$

For linear problems this will preserve energy exactly:

Lemma 4.1 *Assume the system is of the form $\mathbf{f}(\mathbf{x}) = \mathbf{J}_n^{-\Delta} \nabla H(\mathbf{x}) = \mathbf{H}\mathbf{x} + \mathbf{c}$, where \mathbf{H} is a constant Hamiltonian matrix. Then the exponential Euler method (4.2) preserves energy, i.e., $H(\mathbf{x}_+) = H(\mathbf{x})$.*

Proof The local problem now is (see (3.2))

$$\xi'(t) = \mathbf{F} \xi(t) + \mathbf{U}^\dagger (\mathbf{H}\mathbf{x} + \mathbf{c}), \quad \xi(0) = 0.$$

Then $\xi_+ = \xi(h) = h \phi(h\mathbf{F}) \mathbf{U}^\dagger (\mathbf{H}\mathbf{x} + \mathbf{c})$, i.e., the exponential Euler approximation gives the exact solution for the problem in \mathbb{R}^{2k} . Hence the energy is preserved in the small system and consequently also for $\mathbf{x}_+ = \mathbf{x} + \mathbf{U} \xi(h)$. \square

¹ We use the shorthand notation $\mathbf{x} = \mathbf{x}_j$, $\mathbf{x}_+ = \mathbf{x}_{j+1}$ etc. for the rest of the manuscript.

4.2 Explicit exponential midpoint rule

We consider next the explicit exponential midpoint rule (EEMP)

$$\mathbf{x}_+ = \mathbf{x} + e^{h\mathbf{H}}(\mathbf{x}_- - \mathbf{x}) + 2h \phi(h\mathbf{H})\mathbf{f}(\mathbf{x}), \tag{4.3}$$

where $\mathbf{H} = D\mathbf{f}(\mathbf{x})$ (see also [13]). For linear Hamiltonian problems $\mathbf{f}(\mathbf{x}) = \mathbf{H}\mathbf{x} + \mathbf{c}$ this gives

$$\begin{aligned} \mathbf{x}_+ &= \mathbf{x} + e^{h\mathbf{H}}(\mathbf{x}_- - \mathbf{x}) + 2(e^{h\mathbf{H}} - \mathbf{I})\mathbf{x} + 2h \phi(h\mathbf{H})\mathbf{c} \\ &= e^{h\mathbf{H}}(\mathbf{x}_- + \mathbf{x}) - \mathbf{x} + 2h \phi(h\mathbf{H})\mathbf{c}, \end{aligned}$$

i.e., $\frac{1}{2}(\mathbf{x}_+ + \mathbf{x}) = e^{h\mathbf{H}}\frac{1}{2}(\mathbf{x} + \mathbf{x}_-) + h \phi(h\mathbf{H})\mathbf{c} = \tilde{\mathbf{x}}(h)$, where $\tilde{\mathbf{x}}$ is the solution of $\tilde{\mathbf{x}}'(t) = \mathbf{J}_n^{-1} \nabla H(\tilde{\mathbf{x}}(t))$, $\tilde{\mathbf{x}}(0) = \frac{1}{2}(\mathbf{x} + \mathbf{x}_-)$. Hence the energy of the averages is preserved:

$$H(\frac{1}{2}(\mathbf{x}_+ + \mathbf{x})) = H(\frac{1}{2}(\mathbf{x} + \mathbf{x}_-)). \tag{4.4}$$

Remark 4.1 In [9] it was noticed that the explicit midpoint rule (for the homogeneous problem)

$$\mathbf{x}_+ = \mathbf{x}_- + 2h\mathbf{H}\mathbf{x}$$

preserves another quantity: $\omega(\mathbf{H}\mathbf{x}, \mathbf{x}_+) = \omega(\mathbf{H}\mathbf{x}_-, \mathbf{x})$. Equation (4.4) implies this, too.

Again we approximate (4.3) with

$$\mathbf{x}_+ = \mathbf{x} + \mathbf{U} e^{h\mathbf{F}} \mathbf{U}^\dagger(\mathbf{x}_- - \mathbf{x}) + 2h \mathbf{U} \phi(h\mathbf{F}) \mathbf{U}^\dagger\mathbf{f}(\mathbf{x}). \tag{4.5}$$

For this we have the following.

Theorem 4.1 *Let the right hand side \mathbf{f} of the Hamiltonian system be linear, $\mathbf{U} \in \mathbb{R}^{2n \times 2k}$ symplectic, and assume that $\mathbf{x} - \mathbf{x}_-$ is in the range of \mathbf{U} . Then (4.4) holds for the scheme (4.5).*

Proof Now $\mathbf{U}^\dagger \mathbf{U} = \mathbf{I}$. Write $\hat{\mathbf{x}}_+ = \frac{1}{2}(\mathbf{x}_+ + \mathbf{x})$, $\hat{\mathbf{x}} = \frac{1}{2}(\mathbf{x} + \mathbf{x}_-)$. By the assumption there exists $\boldsymbol{\zeta} \in \mathbb{R}^{2k}$ such that $\frac{1}{2}(\mathbf{x} - \mathbf{x}_-) = \mathbf{U}\boldsymbol{\zeta}$. Then $\mathbf{x} = \hat{\mathbf{x}} + \mathbf{U}\boldsymbol{\zeta}$. From (4.5) and $z\phi(z) = e^z - 1$ we get

$$\begin{aligned} \hat{\mathbf{x}}_+ &= \mathbf{x} - \mathbf{U} e^{h\mathbf{F}} \boldsymbol{\zeta} + h \mathbf{U} \phi(h\mathbf{F}) \mathbf{U}^\dagger(\mathbf{H}(\hat{\mathbf{x}} + \mathbf{U}\boldsymbol{\zeta}) + \mathbf{c}) \\ &= \hat{\mathbf{x}} + h \mathbf{U} \phi(h\mathbf{F}) \mathbf{U}^\dagger(\mathbf{H}\hat{\mathbf{x}} + \mathbf{c}) + \mathbf{U} [\mathbf{I} - e^{h\mathbf{F}} + h \phi(h\mathbf{F})\mathbf{F}] \boldsymbol{\zeta} \\ &= \hat{\mathbf{x}} + h \mathbf{U} \phi(h\mathbf{F}) \mathbf{U}^\dagger(\mathbf{H}\hat{\mathbf{x}} + \mathbf{c}). \end{aligned}$$

Thus the $\hat{\mathbf{x}}$ -vectors propagate according to the exponential Euler method (4.2) and we get the result by Lemma 4.1. □

We also have the following.

Lemma 4.2 *Assume that U is a full rank matrix at x with a left inverse U^\dagger , and that $R(U)$ contains $x_- - x$. Then, the approximate explicit exponential midpoint rule is symmetric.*

Proof Multiplying (4.5) with $U e^{-hF} U^\dagger$ gives

$$U e^{-hF} U^\dagger (x_+ - x) = x_- - x + 2h U e^{-hF} \phi(hF) U^\dagger f(x).$$

Since $e^{-z}\phi(z) = \phi(-z)$ we get

$$x_- = x + U e^{-hF} U^\dagger (x_+ - x) - 2h U \phi(-hF) U^\dagger f(x).$$

Thus the steps backward can be taken by replacing h with $-h$. □

4.3 Implicit exponential midpoint rule

As a third exponential integrator, we consider the implicit exponential midpoint rule (IEMP)

$$\begin{aligned} 0 &= e^{hH} (x - \hat{x}) + h \phi(hH) f(\hat{x}), \\ x_+ &= \hat{x} + e^{2hH} (x - \hat{x}) + 2h \phi(2hH) f(\hat{x}) \end{aligned} \quad (4.6)$$

(see [4]). This gives a symmetric method when the linear part H of f is fixed. When H comes from a linearization of a nonlinear Hamiltonian system (2.1), the method is symmetric if $H = Df(\hat{x})$, where \hat{x} satisfies (4.6).

For linear systems of the form $f(x) = Hx + c$, the second equation of (4.6) can be written equivalently as $x_+ = x + 2h \phi(2hH) f(x)$. Then, x_+ propagates according to the exponential Euler method and the energy is preserved in case U is symplectic (Lemma 4.1).

When we apply (4.6) to the local system, the total approximation is symmetric if H is evaluated at the midpoint \hat{x} .

Lemma 4.3 *Assume that U is a full rank matrix with a left inverse U^\dagger . Suppose H is evaluated at \hat{x} , i.e., $H = Df(\hat{x})$, where \hat{x} satisfies (4.6). Consider the approximation*

$$x_+ = x + U \xi_+, \quad (4.7)$$

where ξ is obtained from applying (4.6) to the local system. Then, (4.7) gives a symmetric method.

Proof Applying (4.6) to the local system (3.2) gives

$$\begin{aligned} 0 &= -e^{hF} \xi + h \phi(hF) U^\dagger f(x + U \xi) \\ \xi_+ &= \xi - e^{2hF} \xi + 2h \phi(2hF) U^\dagger f(x + U \xi). \end{aligned} \quad (4.8)$$

We show that (4.8) leads to a symmetric approximation of the full system. Multiplying the upper equation of (4.8) by e^{hF} , and using the relation $e^z\phi(z) = 2\phi(2z) - \phi(z)$ gives

$$-e^{2hF}\xi + 2h\phi(2hF)U^\dagger f(x + U\xi) = h\phi(hF)U^\dagger f(x + U\xi).$$

Combining this and both equations of (4.8) gives

$$\xi_+ = \xi + e^{hF}\xi.$$

Multiplying this from left by U and adding x gives

$$x_+ - \widehat{x} = Ue^{hF}U^\dagger(\widehat{x} - x),$$

where $\widehat{x} = x + U\xi$ and $x_+ = x + U\xi_+$. Replacing here h with $-h$ and multiplying from the left by $Ue^{hF}U^\dagger$ shows the symmetry. □

The IEMP is symmetric if the Jacobian H and the basis U are the same when considering stepping from x to x_+ and vice versa. This is the case if H is evaluated at \widehat{x} , and if U is generated at \widehat{x} using the Krylov subspace methods described in Sect. 5.

Our numerical strategy is to perform one time step h using the exponential Euler method from x to \tilde{x} in order to approximate the midpoint \widehat{x} . Then after evaluating the Jacobian H and forming the basis U at \tilde{x} using Krylov subspace methods, we solve the implicit equation using fixed point iteration and perform the step of size $2h$ to obtain ξ_+ and $x_+ = x + U\xi_+$.

5 Forming the local basis using Krylov subspace methods

We discuss next the approximation of matrix valued ϕ functions using Krylov subspace methods and show how they are naturally connected to the local approximation discussed in Sect. 4.

When matrix A is large but the operation $v \rightarrow Av$ inexpensive, it is reasonable to use Krylov subspace methods. These work in Krylov subspaces

$$K_k(A, v) = \text{span}\{v, Av, A^2v, \dots, A^{k-1}v\}.$$

Then we have $A K_k(A, v) \subset K_{k+1}(A, v)$. The Arnoldi iteration uses the Gram-Schmidt process and produces an orthonormal basis q^1, \dots, q^k for $K_k(A, v)$. Denote $Q_k = [q^1 \dots q^k]$ and $F_k = Q_k^T A Q_k$, which is a Hessenberg matrix.

If the iteration stops, i.e. $A^k v \in K_k(A, v)$, then

- a) $A K_k(A, v) \subset K_k(A, v)$ and $K_j(A, v) = K_k(A, v)$ for all $j \geq k$,
- b) $A Q_k = Q_k F_k$ and for the spectra we have $\Lambda(F_k) \subset \Lambda(A)$,
- c) If $\varphi(z) = \sum_j a_j z^j$ has convergence radius larger than the spectral radius of A and $w \in R(Q_k)$, then

$$\varphi(A)v = Q_k \varphi(F_k) Q_k^T v.$$

The effectivity of Krylov subspace methods is based on the fact that if the component of $A^k \mathbf{v}$ orthogonal to $K_k(\mathbf{A}, \mathbf{v})$ is small, then things are approximately as above and this can happen already for a reasonable size k . This motivates us to consider the approximation

$$\varphi(\mathbf{A})\mathbf{v} \approx \mathbf{Q}_k \varphi(\mathbf{F}_k) \mathbf{Q}_k^T \mathbf{v}$$

(see [6] and [10]). We refer to [12] for a detailed error analysis.

In case \mathbf{Q}_k gives a symplectic basis for $K_k(\mathbf{A}, \mathbf{v})$, one easily verifies that things are as above for $\mathbf{F}_k = \mathbf{Q}_k^\dagger \mathbf{A} \mathbf{Q}_k$ and with \mathbf{Q}_k^T replaced by \mathbf{Q}_k^\dagger . In this case we use the approximation

$$\varphi(\mathbf{A})\mathbf{v} \approx \mathbf{Q}_k \varphi(\mathbf{F}_k) \mathbf{Q}_k^\dagger \mathbf{v}. \quad (5.1)$$

We show next how the Krylov approximation (5.1) is naturally connected to the strategy of applying exponential integrators to the local system (3.2).

5.1 Equivalence of the Krylov and the local system approximations

Consider the local system (3.2) corresponding to the basis \mathbf{U} which gives a symplectic basis for $K_k(\mathbf{H}, \mathbf{f}(\mathbf{x}))$. Recall from Sect. 4 the strategy of solving the local system

$$\xi'(t) = \mathbf{U}^\dagger \mathbf{f}(\mathbf{x} + \mathbf{U}\xi(t))$$

numerically from $\xi(0) = 0$ up to $\xi(h) = \xi_1$ and setting $\mathbf{x}_+ = \mathbf{x} + \mathbf{U}\xi_1$. As shown in Sect. 4.1, applying the exponential Euler method to the local system gives the approximation

$$\mathbf{x}_+ = \mathbf{x} + h\mathbf{U}\phi(h\mathbf{F})\mathbf{U}^\dagger \mathbf{f}(\mathbf{x}).$$

We immediately see from (5.1) that this is the Krylov subspace approximation of the exponential Euler step (4.1).

As shown in Sect. 4.2, applying the exponential explicit midpoint rule to the local system gives

$$\mathbf{x}_+ = \mathbf{x} + \mathbf{U} e^{h\mathbf{F}} \mathbf{U}^\dagger (\mathbf{x}_- - \mathbf{x}) + 2h\mathbf{U}\phi(h\mathbf{F})\mathbf{U}^\dagger \mathbf{f}(\mathbf{x}).$$

This can be seen as a Krylov subspace approximation of the EEMP step (4.3). Here the vector $\mathbf{x}_- - \mathbf{x}$ has to be in the range of \mathbf{U} . This is discussed in Sect. 5.2.1.

Similarly, if we perform a Krylov approximation of the IEMP step (4.6), and denote $\xi = \widehat{\mathbf{x}} - \mathbf{x}$ and $\xi_+ = \mathbf{x}_+ - \mathbf{x}$, we get the small dimensional system (4.8).

In case the basis matrix \mathbf{U} has orthonormal columns, we get the equivalence of the Krylov and local systems approximations by replacing \mathbf{U}^\dagger by \mathbf{U}^T above.

We next consider iterations which produce a symplectic basis for the Krylov subspace $K_k(\mathbf{H}, \mathbf{f}(\mathbf{x}_0))$, where \mathbf{H} is Hamiltonian.

5.2 Symplectic Krylov processes

In order to obtain good local approximations for a Hamiltonian system with linear part \mathbf{H} we would like to have

- a) A symplectic subspace \mathbf{W} with a corresponding basis.
- b) $K_k(\mathbf{H}, \mathbf{f}) \subset \mathbf{W}$ in order to have polynomials of \mathbf{H} applied to \mathbf{f} represented in \mathbf{W} .
We expect this to be worth pursuing for approximations of $\varphi(\mathbf{H})\mathbf{f}$.

Consider first the Krylov subspace corresponding to \mathbf{H} and \mathbf{v} :

$$K_k(\mathbf{H}, \mathbf{v}) = \text{span}\{\mathbf{v}, \mathbf{H}\mathbf{v}, \mathbf{H}^2\mathbf{v}, \dots, \mathbf{H}^{k-1}\mathbf{v}\},$$

and set $W_k = K_k(\mathbf{H}, \mathbf{v}) + \mathbf{J}_n K_k(\mathbf{H}, \mathbf{v})$.

- Now $\mathbf{H} W_k(\mathbf{H}, \mathbf{v}) \not\subset W_{k+1}(\mathbf{H}, \mathbf{v})$, generally.
- If p is a degree $k - 1$ polynomial, then $p(\mathbf{H})\mathbf{v} \in K_k(\mathbf{H}, \mathbf{v}) \subset W_k$.

The construction of a symplectic basis for W_k is slightly more complicated than the standard Arnoldi process. We consider the following three processes.

Symplectic Arnoldi

In the first approach we simply reorthogonalize with respect to $\langle \cdot, \cdot \rangle$ and $\omega(\cdot, \cdot)$ the $\langle \cdot, \cdot \rangle$ —orthogonal vectors provided by the standard Arnoldi. The result is a symplectic and orthonormal matrix.

1. $\mathbf{q} = \mathbf{v} / \|\mathbf{v}\|, \mathbf{Q} = [\mathbf{q}] \mathbf{V} = \mathbf{Q}$
2. for $\ell = 2, \dots, k$ do $\mathbf{r} = \mathbf{H}\mathbf{q}$
 $\mathbf{r} \leftarrow \mathbf{r} - \mathbf{Q}\mathbf{Q}^T\mathbf{r},$
 if $\mathbf{r} \neq 0$, set $\mathbf{q} = \mathbf{r} / \|\mathbf{r}\|, \mathbf{Q} \leftarrow [\mathbf{Q}, \mathbf{q}],$
 $\mathbf{r} = \mathbf{q} - \mathbf{V}\mathbf{V}^T\mathbf{q} - \mathbf{J}_n\mathbf{V}\mathbf{V}^T\mathbf{J}_n^T\mathbf{q},$
 $\mathbf{V} \leftarrow [\mathbf{V}, \mathbf{r} / \|\mathbf{r}\|].$
 else stop.
3. Set $\mathbf{U} = [\mathbf{V} \mathbf{J}_n \mathbf{V}], \mathbf{F} = \mathbf{U}^T \mathbf{H} \mathbf{U}.$

Here the columns of \mathbf{Q} form an orthonormal basis for $K_k(\mathbf{H}, \mathbf{v})$ and those of \mathbf{U} a symplectic basis for W_k .

Remark 5.1 There is a way to construct matrix \mathbf{F} more economically from the computations of step 2. But anyway the reorthogonalization stays the costly part of this approach.

Isotropic Arnoldi

Mehrmann and Watkins [17] suggest the *isotropic Arnoldi process*, which is a direct $\langle \cdot, \cdot \rangle$ and $\omega(\cdot, \cdot)$ —orthogonalization in the Arnoldi process:

1. $\mathbf{q} = \mathbf{v} / \|\mathbf{v}\|, \mathbf{Q} = [\mathbf{q}]$
2. for $\ell = 2, \dots, k$ do $\mathbf{r} = \mathbf{H}\mathbf{q}$
 $\mathbf{r} \leftarrow \mathbf{r} - \mathbf{Q}\mathbf{Q}^T\mathbf{r} - \mathbf{J}_n\mathbf{Q}\mathbf{Q}^T\mathbf{J}_n^T\mathbf{r},$
 if $\mathbf{r} \neq 0$, set $\mathbf{q} = \mathbf{r} / \|\mathbf{r}\|, \mathbf{Q} \leftarrow [\mathbf{Q}, \mathbf{q}],$
 else stop.

3. Set $U = [QJ_nQ], F = U^T H U$.

Here we obtain a symplectic matrix with orthonormal columns. However its range does not necessarily contain the Krylov subspace $K_k(H, v)$. Thus, generally, this iteration does not have the property that $p(H)v$ is in the span of u^1, \dots, u^k for every polynomial p of degree $k - 1$. Since our present aim is to approximate matrix functions, this iteration can be expected to be less effective for our purposes.² We will see this in the numerical tests.

Also there is a possibility of a breakdown: after orthogonalization we may get $r = 0$ without obtaining any useful information about $K_k(H, v)$.

Hamiltonian Lanczos

Benner, Faßbender, and Watkins have several versions of Hamiltonian Lanczos processes (see [2,18]). The following is a promising one from Watkins (Algorithm 4 of [18]):

1. $u^0 = 0, u^1 = v / \|v\|, \beta_0 = 0,$
2. for $j = 0, 1, \dots, k$ do
 - if $j > 0,$ then
 - $x = H u^j, \alpha_j = \langle J_n v^j, x_j \rangle$
 - $u^{j+1} = x - \alpha_j u^j - \beta_{j-1} u_{j-1}$
 - if $j < k,$ then
 - $v^{j+1} = H u^{j+1}, \tau = \langle J_n v^{j+1}, u^{j+1} \rangle$
 - if $\tau \approx 0$ then stop (breakdown).
 - $\sigma = \sqrt{|\tau|}, \delta_{j+1} = \text{sgn } \tau$
 - $u^{j+1} \leftarrow u^{j+1} / \sigma, v^{j+1} \leftarrow \delta_{j+1} v^{j+1} / \sigma$
 - if $j > 0,$ then $\beta_j = \sigma$
3. Form the matrices

$$U = [u^1, \dots, u^k, v^1, \dots, v^k], \quad F = \begin{bmatrix} 0 & T \\ D & 0 \end{bmatrix},$$

where $T = \begin{bmatrix} \alpha_1 & \beta_1 & & & & & \\ \beta_1 & \alpha_2 & \beta_2 & & & & \\ \beta_2 & & \ddots & \ddots & & & \\ & & & \ddots & \beta_{k-1} & & \\ & & & & \beta_{k-1} & \alpha_k & \\ & & & & & & \end{bmatrix}$ and $D = \begin{bmatrix} \delta_1 & & & & & & \\ & \delta_2 & & & & & \\ & & \ddots & & & & \\ & & & \ddots & & & \\ & & & & \delta_k & & \end{bmatrix}.$

Then $U : \mathbb{R}^{2k} \rightarrow \mathbb{R}^{2n}$ is symplectic, its range contains the vectors $H^j v, j = 0, \dots, 2k - 1,$ and $F = U^\dagger H U.$

Due to short recursion this is an economic iteration. But it has similar problems as the usual biorthogonal Lanczos, e.g., near breakdowns and loss of orthogonality. These can be partly circumvented. For small k this may be a good choice.

By the very construction of these symplectic maps we get the following:

Proposition 5.1 *Combining any of the symplectic Krylov processes with a method that preserves energy for the small dimensional system (3.2) will preserve the energy of the original system, too.*

² Mehrmann and Watkins use the iteration for computing eigenpairs of skew-Hamiltonian/Hamiltonian pencils.

The costs of these algorithms are approximately as follows. To produce a basis of dimension k , the Hamiltonian Lanczos requires k matvecs (matrix vector multiplications) and $2k$ inner products, the Arnoldi iteration k matvecs and $k^2/2$ inner products, the Isotropic Arnoldi $k/2$ matvecs and k^2 inner products, and the Symplectic Arnoldi $k/2$ matvecs and $2k^2$ inner products. However, only the Hamiltonian Lanczos and the Arnoldi iteration provide then also a basis for a Krylov subspace of dimension k . Thus the Hamiltonian Lanczos process is the cheapest alternative to obtain a basis for a Krylov subspace of a given dimension.

The main weaknesses of each of these algorithms are

- Arnoldi in \mathbb{R}^{2n} : the approximation is not Hamiltonian.
- Hamiltonian Lanczos: breakdown, early loss of symplecticity.
- Isotropic Arnoldi: does not include a Krylov subspace.
- Symplectic Arnoldi: expensive.

5.2.1 Symplectic reorthogonalization of a vector to the basis

When using the EEMP method (4.3), the vector $\mathbf{x}_- - \mathbf{x}$ needs to be added to the basis U_k at each time step. For orthogonal and/or isotropic basis this is straightforward. For the symplectic basis, a symplectic version of the Gram–Schmidt algorithm adds \mathbf{x} and $J_n \mathbf{x}$ to the basis $U = [V \ W]$. This algorithm is shown in the following pseudocode. Here the symplectic orthogonalization can also be performed in a modified Gram–Schmidt manner, one vector at a time. Notice also that in the second step the vector $\hat{\mathbf{x}}$ could be scaled with any constant.

1. $\hat{\mathbf{x}} = \mathbf{x} - \sum_{\ell=1}^n \omega(\mathbf{w}_\ell, \mathbf{x}) \mathbf{v}_\ell + \sum_{\ell=1}^n \omega(\mathbf{v}_\ell, \mathbf{x}) \mathbf{w}_\ell$
2. $\mathbf{v}_{k+1} \leftarrow \hat{\mathbf{x}}$
3. $\mathbf{x} \leftarrow J_n \hat{\mathbf{x}}$
4. $\tilde{\mathbf{x}} = \mathbf{x} - \sum_{\ell=1}^n \omega(\mathbf{w}_\ell, \mathbf{x}) \mathbf{v}_\ell + \sum_{\ell=1}^n \omega(\mathbf{v}_\ell, \mathbf{x}) \mathbf{w}_\ell$
5. $\mathbf{w}_{k+1} \leftarrow -\frac{\tilde{\mathbf{x}}}{\omega(\mathbf{v}_{k+1}, \tilde{\mathbf{x}})}$

6 Numerical tests

We compare numerically the three exponential time integrators of Sect. 4 and the four Arnoldi like processes of Sect. 5 to produce the local basis U_k . We apply the methods to large sparse Hamiltonian systems which are obtained from finite difference discretizations of one dimensional nonlinear wave equations. For ease of presentation, we first illustrate by an example our approach of deriving large sparse Hamiltonian systems from Hamiltonian PDEs. For further examples we refer to [5].

6.1 Spatial discretization of Hamiltonian PDEs

As an example consider the nonlinear Klein–Gordon equation in one dimension,

$$u_{tt} = u_{xx} - f(u), \tag{6.1}$$

where $u(x, t)$ is a scalar valued periodic function ($u(0, t) = u(L, t)$ for some $L > 0$) and f is a smooth function. To obtain a Hamiltonian system approximating the equation (6.1), we perform a discretization with respect to x on an equidistant grid with an interval $\Delta x = L/n, n \in \mathbb{N}$, and denote by $q_i(t)$ and $p_i(t)$ the approximations to $u(i \Delta x, t)$ and $u_t(i \Delta x, t)$. For the second derivative u_{xx} we use the central difference approximation. Expressing the approximations as vectors $\mathbf{q}(t)$ and $\mathbf{p}(t)$ (i.e. $(\mathbf{q})_i = q_i$ and $(\mathbf{p})_i = p_i$), we get the approximation of the PDE in a matrix form as

$$\mathbf{p}'(t) = \mathbf{\Delta}_n \mathbf{q}(t) - \mathbf{f}(\mathbf{q}(t)),$$

where $\{\mathbf{f}(\mathbf{q})\}_i = f(q_i)$ and $\mathbf{\Delta}_n \in \mathbb{R}^{n \times n}$ is the discretized Laplacian with periodic boundary conditions,

$$\mathbf{\Delta}_n = \frac{1}{(\Delta x)^2} \begin{bmatrix} -2 & 1 & & & 1 \\ 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots \\ & & & 1 & -2 & 1 \\ 1 & & & & 1 & -2 \end{bmatrix}. \tag{6.2}$$

Defining the Hamiltonian function

$$H(\mathbf{q}, \mathbf{p}) = \frac{1}{2} \mathbf{p}^T \mathbf{p} - \frac{1}{2} \mathbf{q}^T \mathbf{\Delta}_n \mathbf{q} + \sum_{i=1}^n F(q_i), \tag{6.3}$$

where $F'(u) = f(u)$, we see that

$$\mathbf{p}'(t) = -\nabla_{\mathbf{q}} H(\mathbf{q}(t), \mathbf{p}(t)).$$

Setting $\mathbf{x}(t) = \begin{bmatrix} \mathbf{q}(t) \\ \mathbf{p}(t) \end{bmatrix}$, we have the Hamiltonian system in \mathbb{R}^{2n}

$$\mathbf{x}'(t) = \mathbf{J}_n^{-1} \nabla H(\mathbf{x}(t)), \quad \mathbf{x}(0) = \mathbf{x}_0, \tag{6.4}$$

where $\mathbf{x}_0 = \begin{bmatrix} \mathbf{q}_0 \\ \mathbf{p}_0 \end{bmatrix}$ comes from the discretization of the initial values of (6.1).

6.2 Linear wave equation

As a first numerical example we consider the linear wave equation with the Dirichlet boundary conditions,

$$\begin{aligned} \partial_t u(x, t) &= \partial_{xx} u(x, t) + g(x) \\ u(x, 0) &= u_0(x), \quad u_t(x, 0) = v_0(x) \\ u(0, t) &= u(L, t) = 0, \end{aligned}$$

where $x \in [0, L], t \in [0, T]$, and

$$g(x) = \frac{1}{8}(x(x - L))^2, \quad u_0(x) = \frac{1}{1 + \sin^2(\pi x)} - 1, \quad v_0(x) = 0.$$

Performing spatial discretization on an equidistant grid of size n using central differences leads to a Hamiltonian system of the form (6.4) with the Hamiltonian (6.3), where $F(\mathbf{q}) = -\mathbf{c}^T \mathbf{q}$, $(\mathbf{c})_i = g(x_i)$. Here the initial data $\mathbf{x}_0 = \begin{bmatrix} q_0 \\ p_0 \end{bmatrix}$, where $(\mathbf{q}_0)_i = u_0(x_i)$, $(\mathbf{p}_0)_i = v_0(x_i)$, $x_i = i \Delta x$, and $\Delta x = L/(n + 1)$. Here the discretized Laplacian is given by

$$\Delta_n = \frac{1}{(\Delta x)^2} \begin{bmatrix} -2 & 1 & & & \\ 1 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & & 1 & -2 \\ & & & & 1 & -2 \end{bmatrix}.$$

We set $L = 2, n = 400$, and we integrate up to $T = 50$ with the step size $h = T/n_t$, where $n_t = 2000$.

Using this linear example we illustrate the differences between the iterative processes of Sect. 5 to produce the basis $\mathbf{U}_k \in \mathbb{R}^{2n \times 2k}$. We apply the exponential Euler method (4.2) to the small dimensional system (3.2) obtained from the projection using \mathbf{U}_k .

As illustrated in Fig. 1, the approximation obtained using the Arnoldi iteration results generally in a linear growth of the energy error, whereas the symplectic basis gives a bounded energy error. Figure 2 shows that, as opposed to the Hamiltonian Lanczos approximation, the energy error of the Arnoldi approximation is dependent on the accuracy of the approximation. Notice that in both cases

$$K_\ell(\mathbf{H}, \mathbf{f}_0) \subset \text{Range}(\mathbf{U}_k), \quad \text{where } \ell = \dim(\mathbf{U}_k), \tag{6.5}$$

and \mathbf{f}_0 is the right hand side of (6.4) evaluated at $x(0)$. Property (6.5) means that these processes give a polynomial approximation of degree ℓ for the exponential Euler step which gives the exact solution at $t = h$. This effect is also seen in Fig. 3, which depicts the solution errors for the Arnoldi iteration and the Hamiltonian Lanczos process. When $\dim(\mathbf{U}_k) = 16$, the methods give errors not far from each other, however for smaller basis size the symplectic alternative gives more accurate results.

When increasing the basis size also the isotropic Arnoldi and the symplectic Arnoldi start to perform better (see Fig. 4). Need for a larger dimension is expected for the symplectic Arnoldi since instead of (6.5), only $K_{\ell/2}(\mathbf{H}, \mathbf{f}_0) \subset \text{Range}(\mathbf{U}_k)$, where $\ell = \dim(\mathbf{U}_k)$. Isotropic Arnoldi performs worse as expected due to its poor polynomial approximation properties. However, both processes give bounded energy errors as in both cases \mathbf{U}_k is symplectic (Fig. 5).

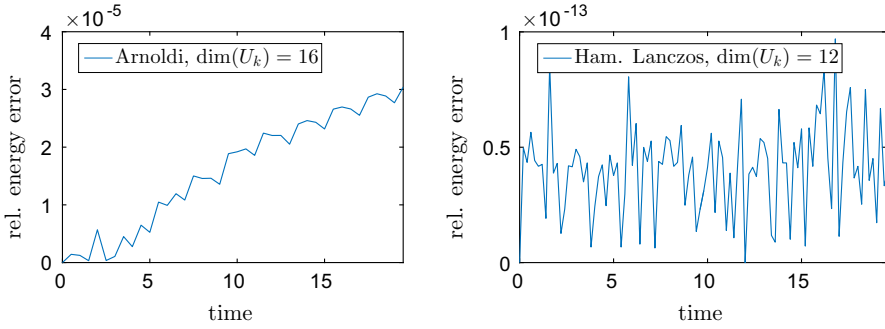


Fig. 1 Linear wave equation and relative energy errors for the EE, when U_k is produced by the Arnoldi iteration ($U_k \in \mathbb{R}^{2n \times 16}$) and the Hamiltonian Lanczos process ($U_k \in \mathbb{R}^{2n \times 12}$)

Fig. 2 Linear wave equation and relative energy errors for the EE, when U_k is produced by the Arnoldi iteration and the Hamiltonian Lanczos process

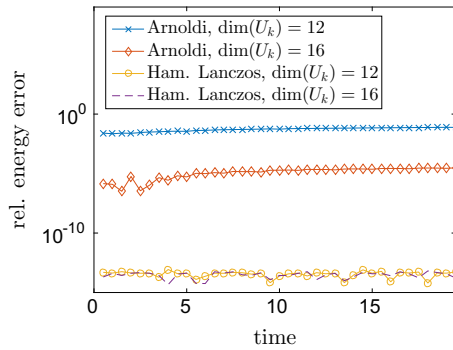
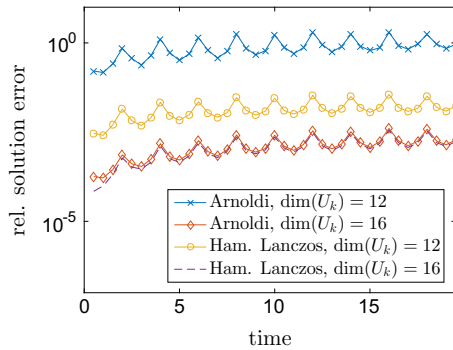


Fig. 3 Linear wave equation and relative solution errors for the EE, when U_k is produced by the Arnoldi iteration and the Hamiltonian Lanczos process



6.3 Nonlinear Schrödinger equation

Consider next a one dimensional nonlinear Schrödinger equation (NLS) on $[-4\pi, 4\pi]$ with periodic boundary conditions (see [3]),

$$\begin{aligned}
 i \partial_t \psi(x, t) &= -\frac{1}{2} \partial_{xx} \psi(x, t) + |\psi(x, t)|^2 \psi(x, t) - V_0 \sin^2(x) \psi(x, t) \\
 \psi(x, 0) &= \psi_0(x), \quad \text{for all } x \in [-4\pi, 4\pi] \\
 \psi(-4\pi, t) &= \psi(4\pi, t).
 \end{aligned}$$

Fig. 4 Linear wave equation and relative solution errors for the EE, when U_k is produced by the symplectic Arnoldi iteration and the isotropic Arnoldi process

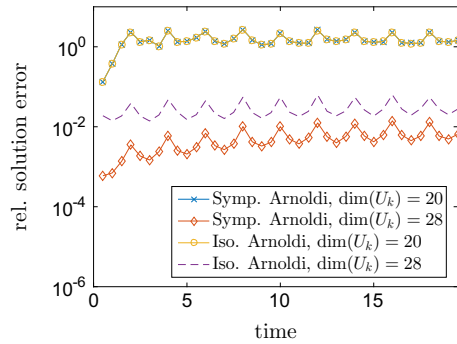
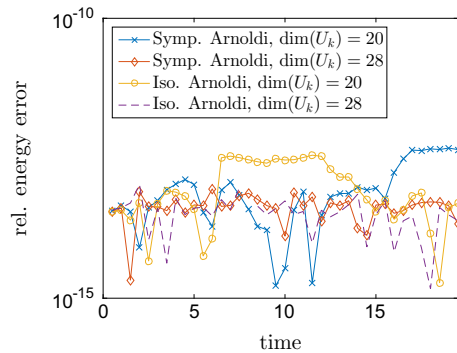


Fig. 5 Linear wave equation and relative energy errors for the EE, when U_k is produced by the symplectic Arnoldi iteration and the isotropic Arnoldi process



The initial value is given by

$$\psi_0(x) = \sqrt{V_0 \sin^2(x) + B} e^{i\theta(x)}, \tag{6.6}$$

where the phase function $\theta(x)$ satisfies

$$\tan(\theta(x)) = \pm \sqrt{1 + V_0/B} \tan(x)$$

(see Fig. 6). We set $V_0 = B = 1.0$ which gives a stable soliton solution (see [3]).

We first carry out a spatial discretization on an equidistant grid with grid size $\Delta x = 8\pi/n$ and denote by $q_i(t)$ and $p_i(t)$ the approximations to $\text{Re } \psi(i \Delta x, t)$ and $\text{Im } \psi(i \Delta x, t)$. This leads to a Hamiltonian system with the energy functional

$$H(\mathbf{p}, \mathbf{q}) = -\frac{1}{4} (\mathbf{q}^T \Delta_n \mathbf{q} + \mathbf{p}^T \Delta_n \mathbf{p}) + \frac{1}{4} \sum_{i=1}^n (q_i^2 + p_i^2)^2 - \frac{V_0}{2} \sum_{i=1}^n \sin^2(x_i) (q_i^2 + p_i^2), \tag{6.7}$$

Fig. 6 The phase function $\theta(x)$ of the initial data (6.6)

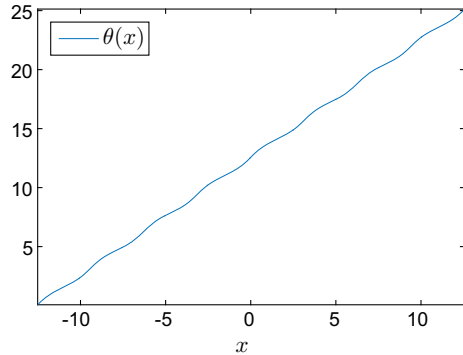
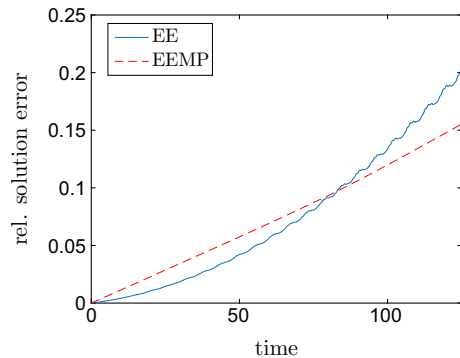


Fig. 7 NLS and relative solution errors for the EE and the EEMP, when U_k is produced by the Arnoldi iteration. Here $\dim(U_k) = 20$



where Δ_n is a discretized Laplacian of the form (6.2). We get from (6.7) the Hamiltonian system

$$\begin{aligned} \begin{bmatrix} q'(t) \\ p'(t) \end{bmatrix} &= -\frac{1}{2}J^{-1} \begin{bmatrix} \Delta_n & 0 \\ 0 & \Delta_n \end{bmatrix} \begin{bmatrix} q(t) \\ p(t) \end{bmatrix} + J^{-1} \begin{bmatrix} (q(t) \circ q(t) + p(t) \circ p(t)) \circ q(t) \\ (q(t) \circ q(t) + p(t) \circ p(t)) \circ p(t) \end{bmatrix} \\ &\quad - J^{-1}V_0 \begin{bmatrix} \sin^2(\bar{x}) \circ q(t) \\ \sin^2(\bar{x}) \circ p(t) \end{bmatrix}, \end{aligned} \tag{6.8}$$

where $\{\bar{x}\}_i = i \Delta x$, \circ denotes the Hadamard product, $(u \circ v)_i = u_i v_i$, and $\sin^2(\bar{x})_i = \sin^2(x_i)$. The initial condition $\begin{bmatrix} q_0 \\ p_0 \end{bmatrix}$ comes from the discretization of $\psi_0(x)$.

We set $n = 500$ and integrate first from 0 to $T = 40\pi$ with the step size $h = T/n_t$, $n_t = 8000$. The benefits obtained from the symmetry properties of the EEMP (Lemma 4.2) are illustrated by Figs. 7 and 8 which depict the relative energy errors and solution errors given by the EE and the EEMP, when $\dim(U_k) = 20$. The nonsymmetric EE shows a linear growth in energy error and quadratic growth in solution error whereas the EEMP gives a bounded energy error and a linear growth in solution error.

We then integrate from 0 to $T = 80\pi$ with the step size $h = T/n_t$, $n_t = 10000$, which implies that the norm of hH is now bigger and thus larger dimension for U_k is required. Differences resulting from the symplecticity of the basis U_k can be seen

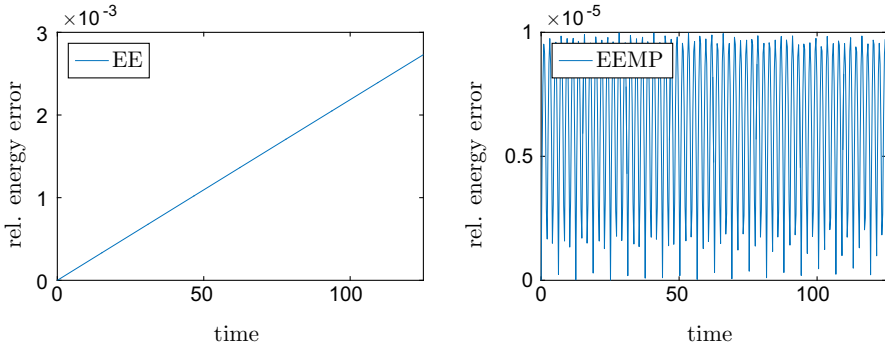
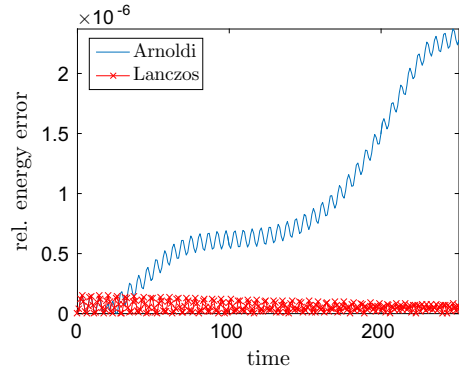


Fig. 8 NLS and relative energy errors for the EE and the EEMP, when U_k is produced by the Arnoldi iteration. Here $\dim(U_k) = 20$

Fig. 9 NLS and relative energy errors for the IEMP when the basis U_k is produced by the Arnoldi and by the Hamiltonian Lanczos iteration. Here $\dim(U_k) = 24$



in Fig. 9 where we compare the IEMP when $\dim(U_k) = 24$ and U_k is given by the Arnoldi iteration and the Hamiltonian Lanczos process. The Arnoldi iteration shows a growth of energy error whereas the Hamiltonian Lanczos iteration shows bounded energy error.

6.4 Nonlinear Klein–Gordon equation

As a last numerical example we consider the nonlinear Klein–Gordon equation with periodic boundary conditions,

$$\begin{aligned} \partial_{tt}u(x, t) &= \partial_{xx}u(x, t) - m^2u(x, t) - gu(x, t)^3 \\ u(x, 0) &= u_0(x), \quad \text{for all } x \in [0, L] \\ u(0, t) &= u(L, t). \end{aligned}$$

After a spatial discretization on the interval $[0, L]$ using the finite differences with the grid size $\Delta x = L/n$, we get a Hamiltonian system of the form (6.4) with the

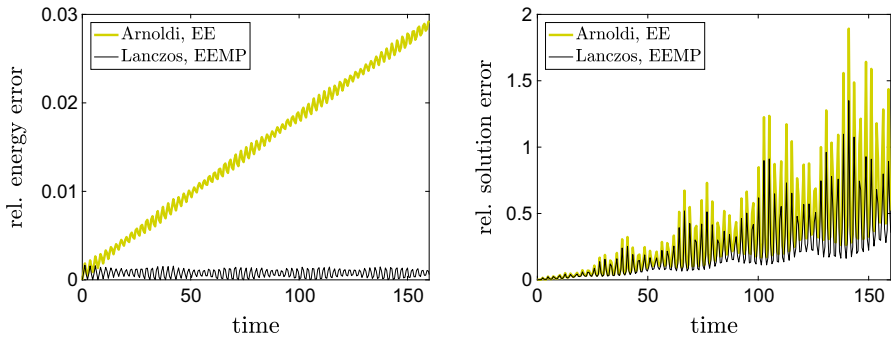
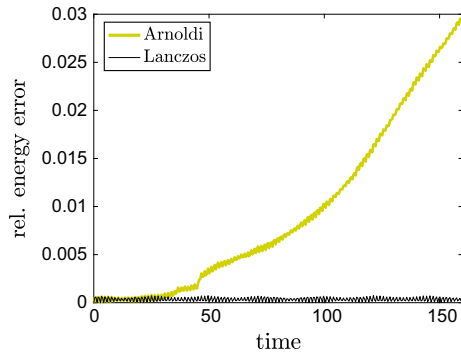


Fig. 10 Klein–Gordon equation and relative energy (left) and solution (right) errors for the EE and the EEMP when the basis $U_k \in \mathbb{R}^{2n \times 20}$ is resp. given by the Arnoldi and by the Hamiltonian Lanczos iteration

Fig. 11 Klein–Gordon equation and relative energy errors for the IEMP, when the basis $U_k \in \mathbb{R}^{2n \times 22}$ is produced using the Arnoldi iteration and the Hamiltonian Lanczos process



Hamiltonian (6.3), where $F(q_i) = \frac{m^2}{2}q_i^2 + \frac{g}{4}q_i^4$. We set the initial data

$$u(x, 0) = A \left(1 + \cos\left(\frac{2\pi}{L}x\right) \right)$$

$$u_t(x, 0) = 0.$$

Set $A = 1, m = 0.5, L = 1, g = 1$ and $n = 400$. We integrate from 0 to $T = 180$ with the step size $h = \frac{T}{n_t}, n_t = 9000$.

Here applying the EEMP combined with the Arnoldi iteration results in an unstable method. However, the Hamiltonian Lanczos process gives a stable alternative. Figure 10 show the relative energy and solution errors for the EEMP combined with the Hamiltonian Lanczos, and for the EE combined with the Arnoldi iteration. For the energy error the EEMP combined with a symplectic basis gives a bounded energy error and also a smaller solution error than the EE combined with the Arnoldi iteration.

When applying the IEMP, the effect of the symplecticity of U_k shows up. Figure 11 shows the relative energy errors when $\dim(U_k) = 22$ and U_k is produced using the Arnoldi iteration and the Hamiltonian Lanczos process.

7 Conclusions and outlook

The theoretical background of this research was the following. By backward error analysis it can be shown that applying a symplectic integrator to an integrable Hamiltonian systems gives a symplectic map $\mathbf{x}_j \rightarrow \mathbf{x}_{j+1}$ and also

$$(\#) \quad \begin{cases} \text{small error in energy uniformly,} \\ \text{error growth linear in } t \end{cases}$$

for exponentially long times (see [11, Ch. X]). Behavior (#) can also be shown for symmetric time integrators when applied to integrable Hamiltonian systems (see [11, Ch. XI]).

Here we have considered exponential integration methods which give symplectic maps when applied to linear Hamiltonian systems. When using the approximations to nonlinear systems the resulting maps are not symplectic, but (#) can anyway be observed numerically. This is especially true when using the exponential explicit midpoint rule

$$\mathbf{x}_{j+1} = \mathbf{x}_j + \mathbf{U} e^{h\mathbf{F}} \mathbf{U}^\dagger (\mathbf{x}_{j-1} - \mathbf{x}_j) + 2h \mathbf{U} \phi(h\mathbf{F}) \mathbf{U}^\dagger \mathbf{f}(\mathbf{x}_j)$$

in such a way that the range of \mathbf{U} contains

$$\text{span}\{\mathbf{x}_{j-1} - \mathbf{x}_j, \mathbf{f}(\mathbf{x}_j), \mathbf{H}\mathbf{f}(\mathbf{x}_j), \dots, \mathbf{H}^{k-1}\mathbf{f}(\mathbf{x}_j)\}.$$

Then the method becomes symmetric. The effect of symplecticity of \mathbf{U} can be seen numerically when applying the implicit exponential midpoint rule (Sect. 4.3) to nonlinear Hamiltonian problems (see e.g. Fig. 11).

The numerical experiments clearly show that both preserving the Hamiltonian structure and the time symmetry are important when applying exponential integrators with Krylov approximations to large scale Hamiltonian systems. The Hamiltonian Lanczos method appears to be the most efficient method to produce a symplectic basis \mathbf{U}_k among those alternatives that give a Krylov subspace of a given dimension. However, further study is needed to find iterations with short ω -orthogonalization recursions that are more efficient and numerically stable for the approximation of the ϕ functions.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Abraham, R., Marsden, J.E.: Foundations of Mechanics, 2nd edn. Benjamin/Cummings Publishing Company, Reading (1978)

2. Benner, P., Fassbender, H.: An implicitly restarted symplectic Lanczos method for the Hamiltonian eigenvalue problem. *Linear Algebra Appl.* **263**, 75–111 (1997)
3. Bronski, J.C., Carr, L.D., Deconinck, B., Kutz, J.N.: Bose–Einstein condensates in standing waves: the cubic nonlinear Schrödinger equation with a periodic potential. *Phys. Rev. Lett.* **86**(8), 1402–1405 (2001)
4. Celledoni, E., Cohen, D., Owren, B.: Symmetric exponential integrators with an application to the cubic Schrödinger equation. *Found. Comput. Math.* **8**, 303–317 (2008)
5. Celledoni, E., Grimm, V., McLachlan, R.I., McLaren, D.I., O’Neale, D., Owren, B., Quispel, G.R.W.: Preserving energy resp. dissipation in numerical PDEs using the “Average Vector Field” method. *J. Comput. Phys.* **231**(20), 6770–6789 (2012)
6. Druskin, V.L., Knizhnerman, L.A.: Two polynomial methods of calculating functions of symmetric matrices. *USSR Comput. Math. Math. Phys.* **29**, 112–121 (1989)
7. Eirola, T.: Krylov integrators for Hamiltonian systems. Workshop on Exponential Integrators, October 20–23 (2004) Innsbruck, Austria. <https://numerical-analysis.uibk.ac.at/images/Conference-Bilder/Workshop-Exponential-Integrators/pdfs/timo.pdf>
8. Eirola, T., Huhtanen, M., von Pfaler, J.: Solution methods for \mathbb{R} -linear problems in C^n . *SIAM J. Matrix Anal. Appl.* **25**(3), 804–828 (2003)
9. Eirola, T., Sanz-Serna, J.M.: Conservation of integrals and symplectic structure in the integration of differential equations by multistep methods. *Numer. Math.* **61**, 281–290 (1992)
10. Gallopoulos, E., Saad, Y.: Efficient solution of parabolic equations by Krylov approximation methods. *SIAM J. Sci. Stat. Comput.* **13**, 1236–1264 (1992)
11. Hairer, E., Lubich, C., Wanner, G.: *Geometric Numerical Integration: Structure Preserving Algorithms for Ordinary Differential Equations*. Springer Series in Computational Mathematics, vol. 31, 2nd edn. Springer, Berlin (2006)
12. Hochbruck, M., Lubich, C.: On Krylov subspace approximations to the matrix exponential operator. *SIAM J. Numer. Anal.* **34**, 1911–1925 (1997)
13. Hochbruck, M., Lubich, C., Selhofer, H.: Exponential integrators for large systems of differential equations. *SIAM J. Sci. Comput.* **19**, 1552–1574 (1998)
14. Hochbruck, M., Ostermann, A.: Exponential integrators. *Acta Numer.* **19**, 209–286 (2010)
15. Koskela, A.: Structure preserving Krylov integrators for Hamiltonian systems. Master’s thesis, Aalto University, Finland (2010)
16. Lubich, C.: A variational splitting integrator for quantum molecular dynamics. *Appl. Numer. Math.* **48**, 355–368 (2008)
17. Mehrmann, V., Watkins, D.: Structure-preserving methods for computing eigenpairs of large sparse skew-Hamiltonian/Hamiltonian pencils. *SIAM J. Sci. Comput.* **22**, 1905–1925 (2001)
18. Watkins, D.S.: On Hamiltonian and symplectic Lanczos processes. *Linear Algebra Appl.* **385**, 23–45 (2004)